

Analyse sémantique des communications de l'atelier-recherche Cirad
« Elevage et pauvreté »
tenu à Montpellier les 11 et 12 septembre 2003

Marie-Claude Deboin

Centre de coopération en recherche agronomique pour le développement (Cirad)
 Direction de l'innovation et de la communication (dic), TA383/05, 34398 Montpellier Cedex 5, France

Sommaire

1. Résumé.....	2
2. Contexte de l'étude	2
3. Objectifs de l'étude	3
4. Méthode	3
4.1. Outils d'analyse	3
4.2. Supports d'analyse	4
4.3. Choix du scénario d'analyse	4
4.4. Niveaux d'analyse	5
5. Résultats.....	5
5.1. Résultats d'analyse textuelle en nombre de documents	5
5.2. Résultats d'analyse textuelle en nombre de mots	9
5.2.1. Univers de référence et références utilisées	9
5.2.2. Rapport d'analyse du scénario	10
5.2.2.1. Economie et commerce	11
5.2.2.2. Agriculture et élevage	12
5.2.2.3. Argent et finance	14
5.2.2.4. Pauvreté et richesse	17
5.2.2.5. Fabrication et production	17
5.2.2.6. Aliments	18
5.2.2.7. Politique et syndicalisme	18
5.2.2.8. Pays	20
6. Synthèse des résultats	23
7. Conclusions.....	24
8. Annexes	26
8.1. Annexe 1 : Liste des fichiers et titres des documents analysés	26
8.2. Annexe 2 : Liste des figures	27
8.3. Annexe 3 : Liste des tableaux	27

1. Résumé

Une analyse textuelle des 17 communications rédigées à l'occasion de l'atelier-recherche Cirad « Elevage et pauvreté » qui s'est tenu à Montpellier en 2003, a été menée à l'aide des logiciels d'indexation automatique, de recherche et d'analyse sémantique *Tropes Zoom* de la société *Acetic* (France).

Cette étude, basée sur des classifications thématiques automatiques selon un scénario sémantique prédéfini et sur des analyses, statistique, probabiliste et cognitivo-discursive, a permis d'identifier les grandes variations sémantiques entre documents d'une part, et au sein des textes composant le corpus d'autre part.

Les résultats obtenus montrent que si l'élevage, la pauvreté, la politique, le commerce, les animaux et les aliments sont des thèmes génériques communs et partagés par l'ensemble des textes étudiés, certains champs ou objets d'étude spécifiques à ces thèmes sont rarement ou faiblement abordés par les auteurs. La macro-économie, comme la micro-économie, les notions d'inflation et de déflation, d'exportation, de marketing et d'entrepreneuriat sont ainsi rarement représentées, de même que la politique économique et la géopolitique.

La pauvreté, bien qu'apparaissant être une des préoccupations de l'atelier ne donne pas lieu à des discussions sur des solutions et sur les moyens de résoudre le problème de paupérisation des éleveurs.

Le domaine des technologies se concentre sur les techniques aux dépens de la méthodologie.

La communication se fait par le biais de dialogues ou de questions-réponses sans argumentation.

Enfin, les zones géographiques d'étude concernent essentiellement l'Afrique, notamment l'Afrique de l'Ouest et l'Afrique Centrale, et l'Amérique du Sud, surtout le Brésil et le Pérou. L'Asie, le Proche-Orient, l'Océanie sont des régions peu évoquées.

2. Contexte de l'étude

Au sein du département Elevage et médecine vétérinaire (Emvt) du Centre de coopération internationale en recherche agronomique pour le développement (Cirad), un collectif de chercheurs a été mis en place en 1999 pour réfléchir aux problèmes majeurs et identifier les priorités de recherche sur l'élevage en vue d'accompagner les changements auxquels fait face l'élevage dans les pays en développement.

Ce collectif, baptisé « Economie et Politiques d'Elevage », qui associe des chercheurs d'autres départements du Cirad avec des compétences en socio-économie de l'élevage, organise une formation en économie de l'élevage du 3 au 10 juillet 2004.

La formation, destinée aux chercheurs du Cirad travaillant sur le thème Economie de l'élevage, a pour but d'approfondir les connaissances du groupe sur les méthodologies et les concepts relatifs à l'économie et aux politiques de l'élevage afin de définir les stratégies scientifiques futures du Cirad en relation avec les défis actuels dans le domaine.

Les objectifs spécifiques de l'atelier-formation sont :

- la formation des participants aux concepts de base en économie de l'élevage,
- le référencement de la littérature scientifique sur le thème de l'économie et des politiques d'élevage à partir de l'élaboration d'une base de données bibliographique commune,
- l'identification de thèmes de recherche émergents et particulièrement l'identification de thèmes pour lesquels l'équipe présente un avantage comparatif,
- le montage de nouveaux projets en partenariat avec les pays du Sud.

L'identification de thèmes émergents s'appuiera sur l'étude des documents (ouvrages, manuels, actes d'atelier) issus de la littérature internationale ou produits par le Cirad-emvt dans le domaine de l'économie de l'élevage. Elle se basera également sur l'exploration de la base de données bibliographiques élaborée par l'équipe et sur une analyse bibliométrique de la base.

Enfin, un rapport d'analyse sémantique et statistique sera fourni aux participants, élaboré à partir de l'étude d'un corpus de documents ayant pour thème l'élevage et la pauvreté.

3. Objectifs de l'étude

Afin d'accompagner les travaux de l'atelier et faciliter l'identification des thèmes émergents pour la définition des priorités de recherche par le groupe EPE, une analyse sémantique est proposée aux participants sur la base des communications présentées lors d'un atelier de recherche sur « Elevage et pauvreté » tenu à Montpellier les 11 et 12 septembre 2003.¹ Cette analyse de contenus tente de mettre en évidence les variations sémantiques significatives au sein du corpus des communications présentées dans le cadre du séminaire « Elevage et pauvreté » (cf. annexe 1 : liste des fichiers et titres des communications analysées).

4. Méthode

4.1. Outils d'analyse

La présente étude propose une aide à l'interprétation de textes par l'analyse sémantique et statistique de contenus au moyen de deux outils logiciels :

- *Zoom*, moteur d'indexation automatique et de recherche sémantique,
- *Tropes*, logiciel d'analyse de textes.

Les logiciels *Tropes* et *Zoom*, spécialement conçus par la société française Acetic², permettent :

- d'accélérer la vitesse de lecture de textes,
- d'analyser en profondeur des contenus textuels de façon objective,
- d'extraire l'information pertinente des contenus analysés,
- de classer automatiquement et de structurer rapidement l'information extraite.

Ils s'appuient sur une technologie linguistique utilisant des classifications sémantiques prédéfinies et des méthodes d'analyse statistiques.

Ils permettent d'élaborer des rapports graphiques par l'intermédiaire de l'application *Excel* de *Microsoft*.

Alors que *Zoom* permet de rechercher les concepts sémantiques partagés par un certain **nombre de documents**, *Tropes* élabore des statistiques sur la fréquence globale d'apparition des grandes **catégories de mots** et de leurs sous catégories dans les textes étudiés. *Tropes* fournit également des statistiques sur la cooccurrence et le taux de liaison des classes d'équivalents (mots d'une même famille sémantique) et des catégories de mots.

¹ Elevage et pauvreté : actes de l'atelier-recherche CIRAD, Montpellier, les 11 et 12 septembre 2003. Montpellier : CIRAD, 2003. – [180 p.]

² Acetic, 45, rue Saint Sébastien, 75011 Paris, France, <http://www.acetic.fr/>

Il permet enfin de détecter des propositions remarquables obtenues par contraction du texte et qui introduisent des thèmes ou des personnages principaux exprimant des événements nécessaires à la progression de l'histoire.

Dans la présente étude, les deux applications, *Zoom* et *Tropes*, seront successivement utilisées pour exprimer respectivement les variations sémantiques en nombre de documents et les variations en nombre de mots.

4.2. Supports d'analyse

L'analyse a porté sur 17 textes issus des communications présentées lors de l'atelier et disponibles en format Word (fichiers .doc).

Le corpus des 17 documents soumis à l'analyse correspond à un ensemble de 88995 mots traités par *Tropes Zoom*.

Tous les textes étudiés sont en langue française.

La structure générale, voisine d'un texte à l'autre, inclut : le titre de la communication, l'auteur et son adresse, le résumé introduit par le libellé « Résumé », le corps de la communication, une conclusion introduite par le libellé « Conclusion » et une bibliographie introduite par la mention « Bibliographie ».

Des tableaux et des notes de bas de page sont présents dans certaines communications.

4.3. Choix du scénario d'analyse

Un scénario sémantique est une classification thématique élaborée.

L'analyse de texte avec *Tropes Zoom* s'appuie sur l'analyse grammaticale et sémantique du contenu. Elle procède en étapes successives :

- découpage de phrases en propositions à l'aide de grammaire qui constitue le noyau linguistique des logiciels *Tropes Zoom*,
- levée d'ambiguïté grammaticale des mots du texte,
- élaboration de statistiques,
- identification de classes d'équivalents, mots sémantiquement voisins,
- affichage des résultats d'analyse avec création automatique de rapports graphiques d'analyse.

Plusieurs scénarios de langue française et de langue anglaise sont accessibles par *Tropes Zoom*.

Le scénario en français *Concepts FR V6 – détaillés.scn* a été retenu pour l'analyse des contenus du séminaire « Elevage et pauvreté ».

En raison du niveau d'analyse de la présente étude limitée au nombre de documents qui partagent les mêmes concepts sémantiques, aucune modification - ni ajout, ni suppression, ni correction de termes - n'a été apportée au scénario de base *Concepts FR V6 – détaillés.scn*.

L'indexation des documents selon le scénario retenu, a porté sur l'ensemble des 17 communications disponibles.

Aucun élément des différents textes n'a été écarté de l'analyse par *Zoom Tropes*.

4.4. Niveaux d'analyse

L'analyse effectuée par *Tropes Zoom* peut s'effectuer à différents niveaux du scénario sémantique, c'est-à-dire de la classification thématique des termes. Un rapport d'analyse généré par le logiciel peut ainsi prendre en compte :

- la totalité du scénario, c'est-à-dire tous les concepts identifiés dans les documents aboutissant à une analyse sommaire de contenu,
- toutes les données d'une branche ou d'une classe de scénario (par exemple, uniquement la branche « Economie et commerce »),
- une sélection d'une partie de branche ou de classe de scénario (par exemple, la sous-classe « Import-export » de la branche « Economie et commerce »).

La grande variété de concepts présents dans le corpus de textes analysés nous a conduit à procéder à un examen branche par branche des concepts les plus partagés par les documents, et en leur sein, des variations entre les classes fortement ou faiblement représentées.

5. Résultats

5.1. Résultats d'analyse textuelle en nombre de documents

A partir de l'indexation par *Zoom* du corpus de documents selon le scénario retenu, des rapports d'analyse statistique ont été générés automatiquement sur la base des principaux concepts sémantiques partagés par tout ou partie du corpus des 17 documents étudiés. Les résultats sont exprimés en **nombre de documents** représentant un concept donné.

Zoom permet d'identifier et d'afficher de façon automatique et objective les concepts les plus fréquents traités dans l'ensemble des 17 documents étudiés.

A l'intérieur des classes de concepts résumant les documents analysés, *Zoom* met en évidence les fluctuations dans la représentation de sous-classes sémantiques.

Au sein d'une branche du scénario sémantique fortement représentée dans le corpus de documents analysés, il est ainsi possible d'identifier les thèmes faiblement évoqués et *a contrario*, les thèmes communs partagés par l'ensemble des documents.

Le tableau 1 montre les fluctuations du nombre de documents partageant des classes sémantiques identiques. Seules les classes les plus fréquentes ont été reproduites et au sein de chacune d'elles, seules les fluctuations importantes ont été relevées.

Tableau 1 : Classes sémantiques présentant, en leur sein, des variations importantes de thèmes, exprimées en nombre de documents concernés

Tout le monde parle de ... (Nombre de documents trouvés/ Nombre total)	Peu d'auteurs parlent de ...
Agriculture et élevage (17/17) Agriculture (17/17) Elevage (16/17)	Agriculture biologique (1/17) Agronomie (3/17) Matériel agricole (3/17)
Dans Agriculture (17/17) : Agriculteur (14/17) Ferme (10/17) Récolte (10/17)	Arboriculture (1/17) Horticulture (1/17) Riziculture (1/17)
Dans Elevage (16/17) : Eleveur (13/16) Berger (11/16)	Vacher (1/16) Vivier (1/16) Etable (2/16)
Aliments (17/17) Lait (14/17) Produit laitier (9/17) Viande (12/17)	Miel (1/17) Œuf (1/17) Farine (1/17) Beurre (3/17) Fromage (2/17) Fruit de mer (2/17) Sucre (2/17)
Animaux (17/17) Mammifères (17/17)	Oiseaux (2/17) Insectes (3/17) Poissons (4/17)
Parmi les Mammifères (17/17) : Bovidé (17/17) Ovin (8/17)	Rongeur (1/17) Canidé (2/17) Chien (2/17) Cheval (2/17) Porc (3/17) Camélidé (3/17) Cervidé (3/17)
Commerce (17/17) Laiterie (11/17)	Meunerie (1/17) Fromager (4/17)
Economie et commerce (17/17) Economie et consommation (17/17) Import-export (12/17) Marketing et vente (17/17) Marché et commerce (17/17)	
Dans Economie et consommation Consommation (14)	Inflation et déflation (6)
Dans Import-Export Importation (12/17)	Exportation (7/17)
Dans Marketing et vente (17/17) Vente (16/17)	Marketing (1/17)

Tout le monde parle de ... (Nombre de documents trouvés/ Nombre total)	Peu d'auteurs parlent de ...
Dans Marché et commerce (17/17) Marché (15/17) Concurrence (11/17)	Entrepreneurship (1/17) Fournisseur (1/17)
Argent et Finance (17/17) Fonds et épargne (15/17) Fiscalité (15/17) Moyens de paiement (15/17) Finance (15/17)	Gaspillage (1/17) Comptabilité (2/17)
Dans Finance (15/17) Dépenses (12/17) Crédits (11/17)	Change (2/17) Actionnaire (1/17)
Médecine et santé (17/17) Santé (12/17) Maladies (10/17)	Diététique (2/17) Pharmacie et pharmacologie (1/17)
Ville et urbanisme (17/17) Villes (17/17) Village (10/17)	
Dans Villes (17/17) Ville (14/17) Urbain (12/17)	Bidonville (3/17) Chef-lieu (1/17)
Transports (16/17) Transports terrestres (10/17)	Transports par rail (3/17)
Végétaux (15/17) Céréales (11/17)	Fleurs (2/17) Fruits (5/17)
Alimentation et agro-alimentaire (12/17) Alimentaire (10/12)	Industrie agro-alimentaire (1/17) Agro-alimentaire (2/17)
Pays (17/17) Afrique (16/17) Europe (16/17) dont France (16/17)	Océanie (1/17) Amérique (2/17) Amérique centrale (2/17) Amérique du Sud (3/17) Océan indien (2/17)
Dans Afrique (16/17) Afrique centrale (10/17) Afrique de l'Ouest (11/17)	Afrique du Nord (4/17) Afrique de l'Est (6/17) Afrique méditerranéenne (3/17)
Politique et syndicalisme (16/17) Politique (14/17)	Citoyen (1/17) Syndicalisme (1/17)
Dans Politique (13/17) Politique (13/17)	Géopolitique (2/17)
Technologies et techniques (16/17) Techniques (15/17)	Ingénierie (2/17) Laboratoire (3/17)
Dans Techniques (15/17)	

Tout le monde parle de ... (Nombre de documents trouvés/ Nombre total)	Peu d'auteurs parlent de ...
Technique (9/17)	Méthodologie (4/17)
Communication (16/17) Conversation et dialogue (13/17) Questions et réponses (11/17) Information (10/17)	Message (1/10) Forum (2/17)
Dans Conversation et dialogue (13/17) Débat (11/17)	Argumentation (2/17)
Cognition et connaissance (17/17) Stratégie (14/17) Raisonnement et hypothèse (14/17)	Compétence (2/17) Illettrisme (2/17) Ignorance (3/17)
Cause et effet (16/17) Effet (15/17)	Cause (7/17)
Changement et réorganisation (17/17) Changement (17/17)	Abolition (1/17) Eradication (1/17)

Conformément au contexte de l'atelier « Elevage et pauvreté » dans lequel s'inscrivaient les communications présentées, les 17 documents étudiés partagent les concepts génériques tels que : agriculture et élevage, animaux, aliments, commerce, économie et commerce, argent et finance, médecine et santé, ville et urbanisme, pays.

Les classes, cognition et connaissance ainsi que changement et réorganisation, regroupent également des notions couvertes par l'ensemble des 17 communications.

Au sein de chacune des classes fortement représentées, des différences thématiques significatives sont toutefois observées entre documents.

Ces fluctuations se retrouvent aussi bien dans les objets ou des champs d'étude comme l'élevage, la politique, l'économie, la santé, que dans des processus telles que les connaissances, la communication ou le changement.

C'est ainsi qu'en élevage, il est surtout question d'éleveur ou de berger, rarement de vacher ou d'étable.

Les animaux étudiés sont communément des bovidés. Les ovins sont moins évoqués, les camélidés, les canidés le porc ou le cheval sont des sujets d'étude peu partagés.

Parmi les aliments, les œufs, le beurre ou le fromage ne font pas l'objet d'études communes.

Le secteur de la médecine et santé se réduit à des questions de santé et de maladies. Peu de documents parlent de diététique ou de pharmacologie.

Si tous les auteurs évoquent le marché, le commerce et la consommation, la moitié seulement évoque les problèmes d'inflation et de déflation, d'exportation. Très rares sont ceux qui introduisent les notions de marketing, d'entrepreneuriat ou de fournisseur dans leur écrit.

La politique est également un sujet commun à la majorité des communications. L'approche et l'exploration de ce domaine sont néanmoins peu diversifiées. Les questions de politique économique ou de géopolitique sont quasiment absentes des débats.

Le processus de communication est abordé dans tous les écrits au travers de dialogues, de questions réponses ou d'information. Par contre, l'argumentation n'est pas une approche commune à la majorité des auteurs.

Enfin, si les effets d'un phénomène sont toujours décrits, leurs causes sont plus rarement débattues.

Une analyse complémentaire plus fine axée sur le nombre de mots reliés sémantiquement aux concepts traités dans les documents analysés, peut permettre de préciser les variations sémantiques observées au sein du corpus.

Cette analyse textuelle, exprimée non plus en nombre de documents abordant un concept donné mais en nombre de mots se rattachant à un concept, constitue la seconde étape de notre étude sémantique.

5.2. Résultats d'analyse textuelle en nombre de mots

Le corpus des 17 documents étudiés a été ensuite soumis au logiciel *Tropes* pour une analyse sémantique et statistique de tous les mots des textes selon le scénario original *Concepts FR V6 – détaillés.scn*.

88995 mots ont été traités par *Tropes*. Après être regroupés en plusieurs grandes catégories lexicales par le logiciel, les mots sont affectés dans ces catégories, leur répartition en sous catégories est analysée, ainsi que leur ordre d'arrivée à l'intérieur des propositions et au sein des textes eux-mêmes.

5.2.1. Univers de référence et références utilisées

Les **univers de référence** identifiés par *Tropes* représentent le contexte global d'un texte ou d'un discours.

Ils regroupent, dans des classes d'équivalents, les principales références du texte appelées *références utilisées*.

Les **références utilisées** regroupent elles-mêmes, dans des *classes d'équivalents*, les noms communs et les noms propres ayant un sens voisin et apparaissant fréquemment dans le texte.

Par exemple, « pauvreté », « appauvrissement » et « paupérisation » sont regroupés dans la classe « pauvreté » par *Tropes*. Les mots « troupeau », « troupeaux » et « cheptel » sont regroupés dans la classe « troupeau ».

L'appellation **classes d'équivalents** est utilisée indifféremment par *Tropes Zoom* pour parler des *Univers de référence*, des *Références utilisées* ou des *Groupes sémantiques* du *Scénario*.

Le logiciel génère automatiquement des classes d'équivalents pour tous les mots non référencés dans ses dictionnaires. Ces classes générées contiennent également des mots identifiés par le logiciel mais qui ne peuvent être regroupés par équivalents.

Dans la présente étude, pour des raisons de simplicité d'exploitation, les mots non reconnus par le scénario n'ont pas été intégrés au Scénario original qui n'a pas été ainsi enrichi.

Par exemple, le terme « cirad », inconnu du Scénario *Concepts FR V6 – détaillés.scn*, n'a pas été introduit dans sa classification.

L'analyse par *Tropes* a identifié ce terme et a généré une nouvelle classe « cirad » qui n'apparaît pas dans le Scénario.

Les classes non reconnues ne sont visibles qu'au niveau des *Références utilisées*.

Tropes décrit le contexte d'un discours selon deux niveaux de représentation ;

- l'univers de référence 1 qui représente globalement le contexte,
- l'univers de référence 2 qui propose une représentation plus fine du contexte.

Le tableau 2 cite, parmi les univers de référence et les références les plus significatives identifiées par *Tropes*, les mots les plus fréquents et leur occurrence dans le corpus des textes analysés.

Tableau 2 : Univers de référence 1, univers de référence 2 et références utilisées les plus fréquentes

Nombre de mots	Univers de référence 1	Nombre de mots	Univers de référence 2	Nombre de mots	Références utilisées
1162	agriculture	706	élevage	414	bovidé
905	animal	483	temps	397	texte en anglais
805	commerce	469	mammifère	388	pauvreté
630	afrique	447	production	303	élevage
582	temps	446	lieu	295	éleveur
535	finance	428	aliment	237	paiement
502	richesse	414	agriculture	237	année
500	lieu	410	pauvreté	225	troupeau
491	quantité	404	animal	220	famille
447	production	397	texte en anglais	218	production
430	alimentation	343	vente	200	agriculture
260	famille	325	quantité	181	changement
257	croissance	277	commerce	180	activité
247	rémunération	237	paiement	174	animal
217	sécurité	224	famille	169	prix
216	politique	206	nation	168	exploitation
209	nation	161	système	152	système
201	communication	156	afrique centrale	135	développement
197	économie	147	moyen de paiement	127	ressource
185	cognition	146	afrique de l ouest	125	2002
173	organisation	142	africain	125	producteur
161	système	140	politique	124	nation
151	groupe social	135	développement	122	zone
149	santé	133	insécurité	120	risque

5.2.2. Rapport d'analyse du scénario

Le générateur d'état de *Tropes* permet de construire automatiquement des rapports et tableaux statistiques sous *Excel*, à partir de l'analyse sémantique des textes selon le scénario utilisé.

L'examen de branches du scénario permet d'observer au sein d'une catégorie donnée, les variations d'occurrences des mots dans les différentes classes d'équivalents.

Les figures suivantes visualisent les variations les plus importantes relevées parmi les classes d'équivalents d'occurrence élevée.

5.2.2.1. Economie et commerce

Figure 1 : Répartition des mots dans la classe Economie et commerce (1275 mots)

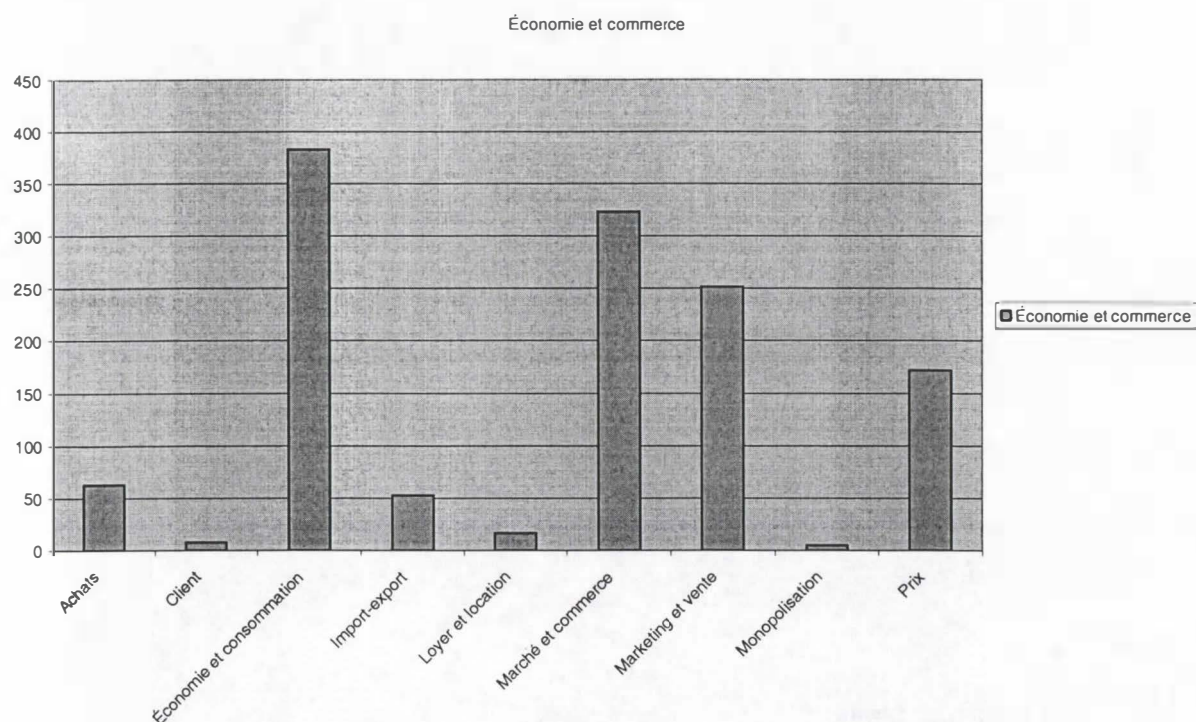


Figure 2 : Répartition des mots dans la sous-classe Marché et commerce (323 mots)

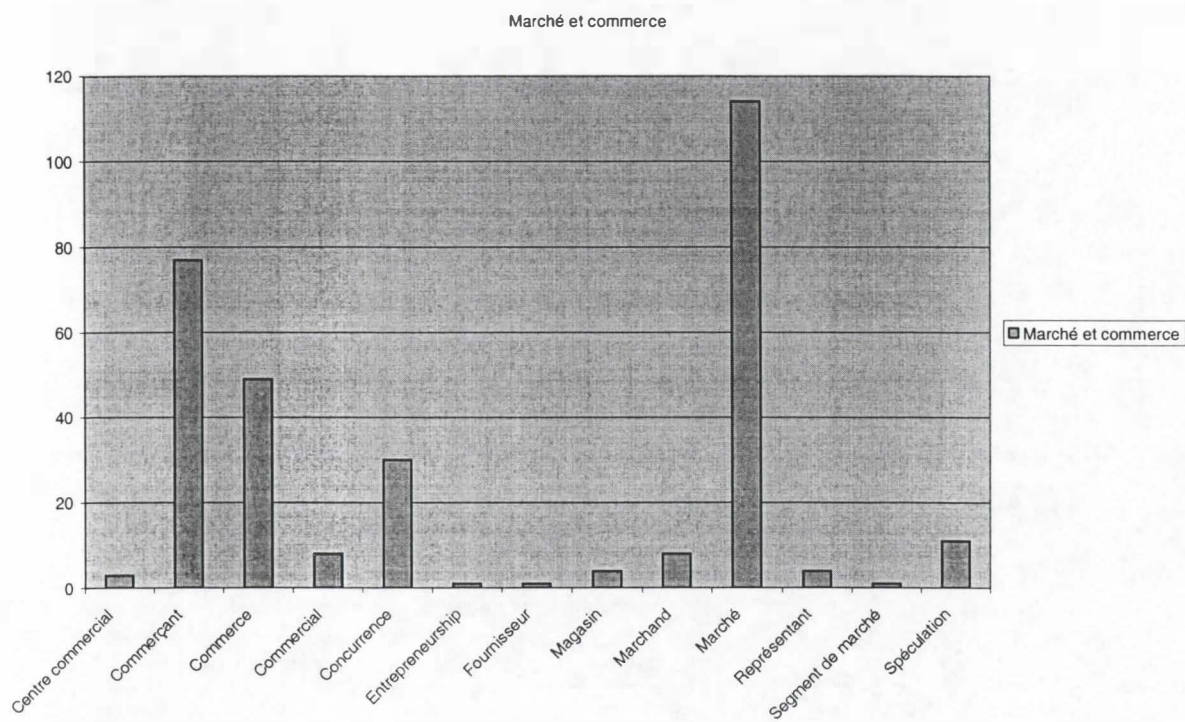
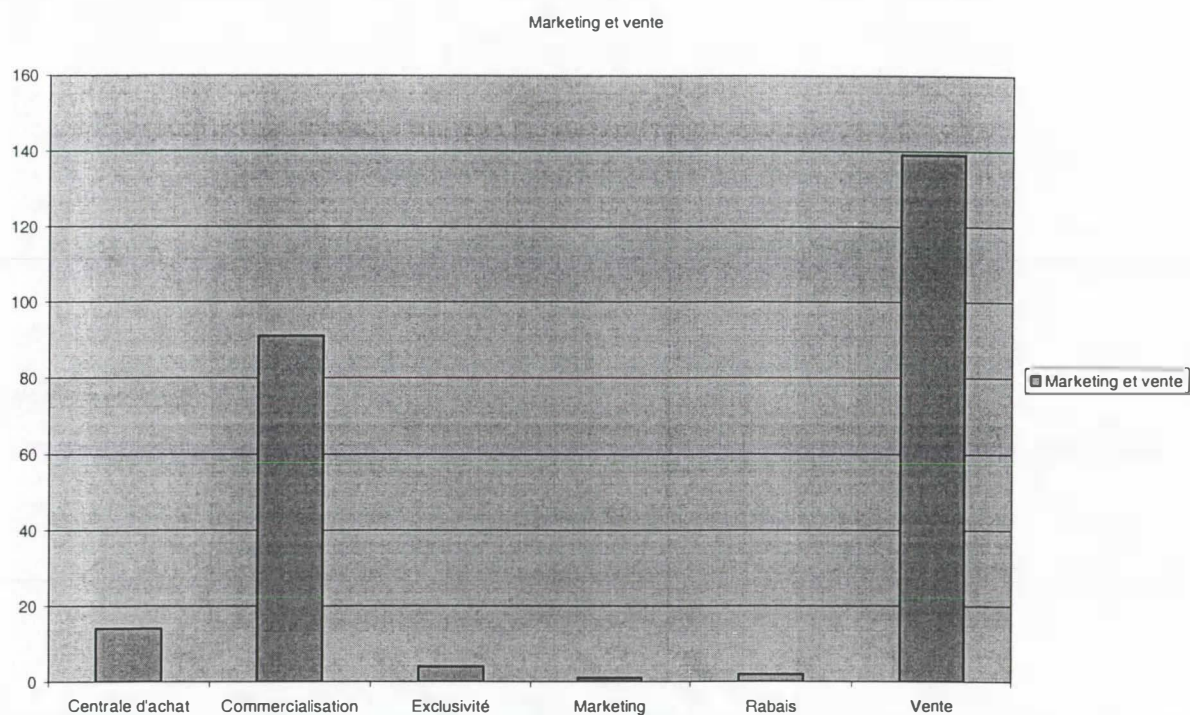


Figure 3 : Répartition des mots dans la sous-classe Marketing et vente (251 mots)

5.2.2.2. Agriculture et élevage

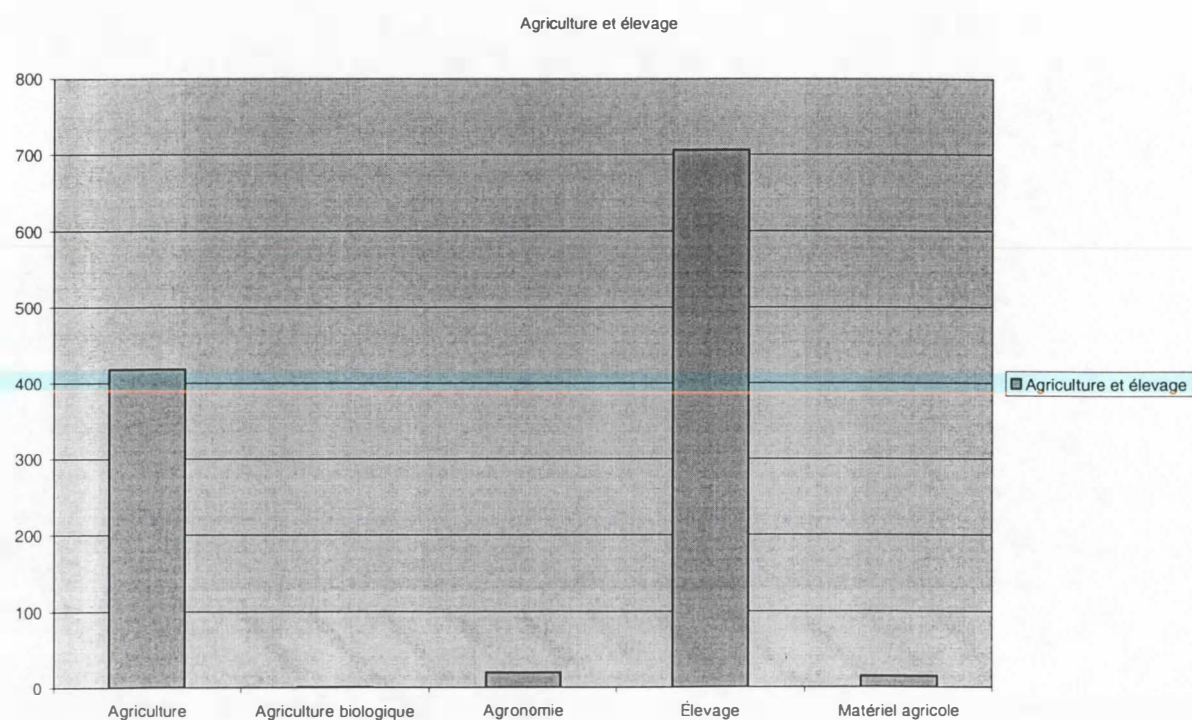
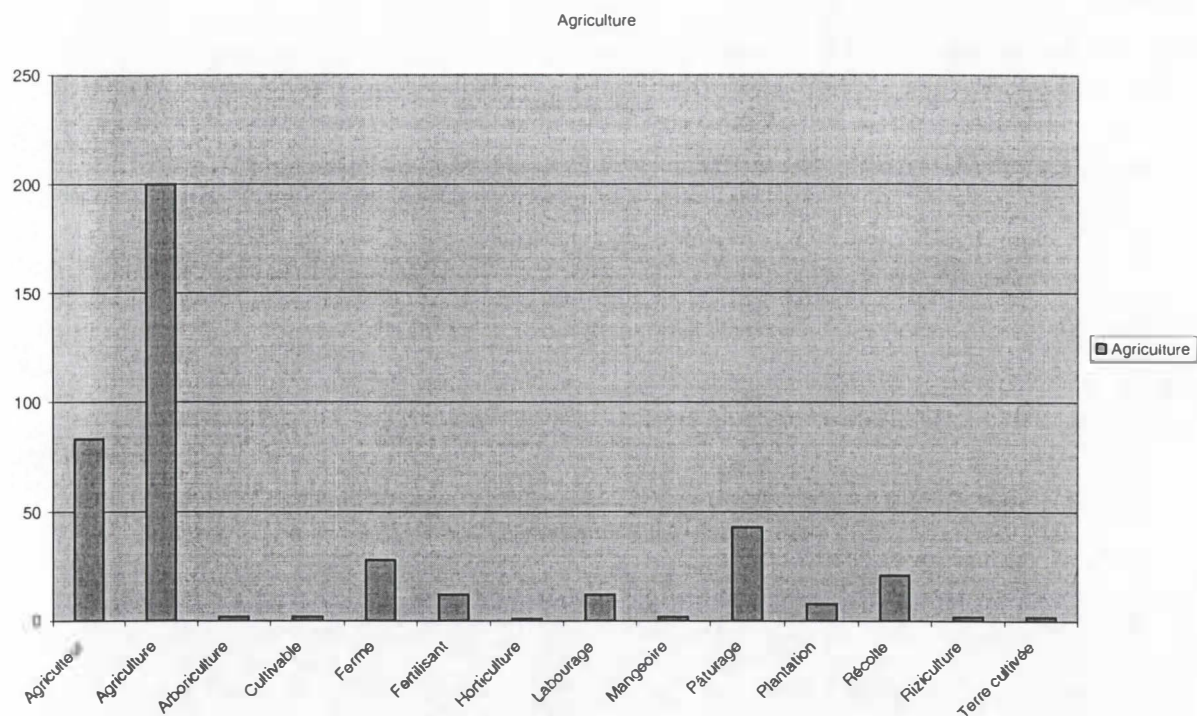
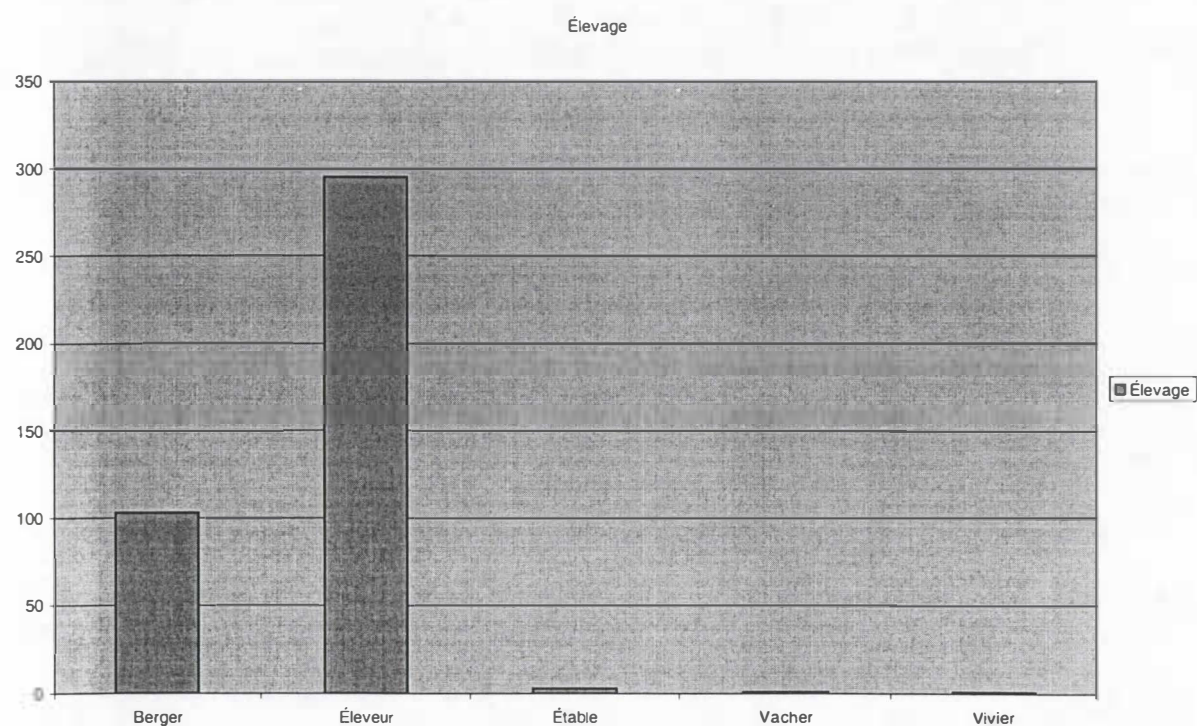
Figure 4 : Répartition des mots dans la classe Agriculture et élevage (1265 mots)

Figure 5 : Répartition des mots dans la sous-classe Agriculture (418 mots)**Figure 6 : Répartition des mots dans la sous-classe Elevage (706 mots)**

5.2.2.3. Argent et finance

Figure 7 : Répartition des mots dans la classe Argent et finance (627 mots)

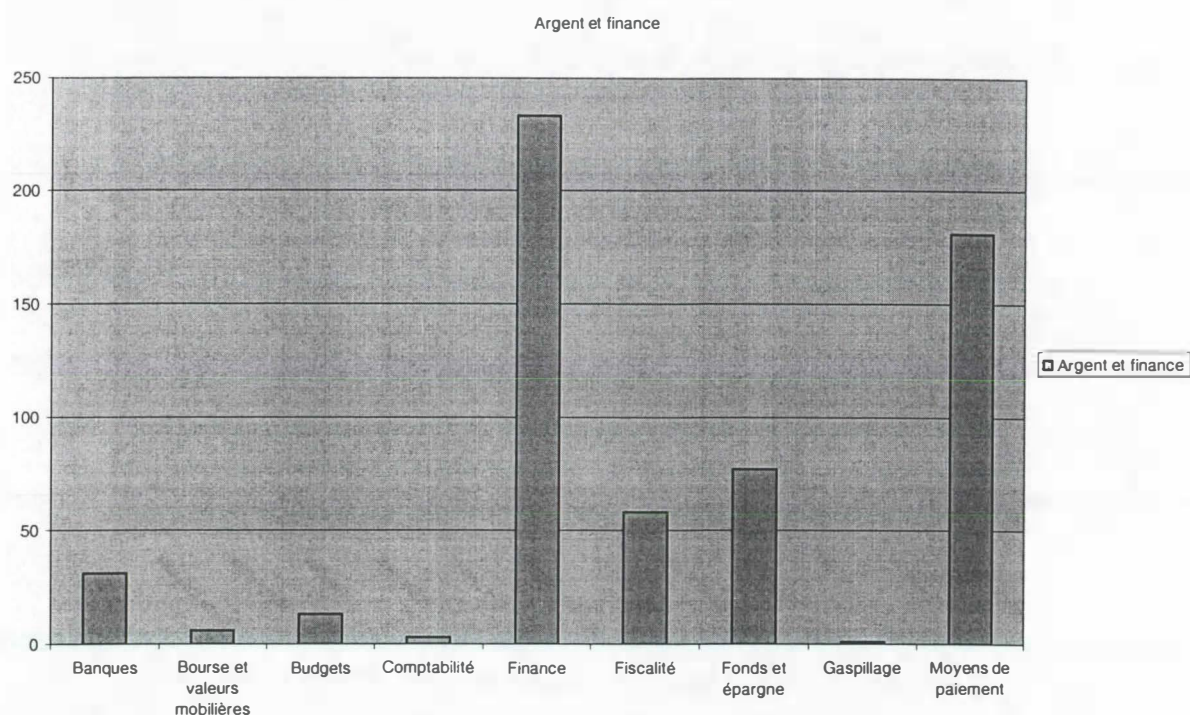


Figure 8 : Répartition des mots dans la sous-classe Finance (233 mots)

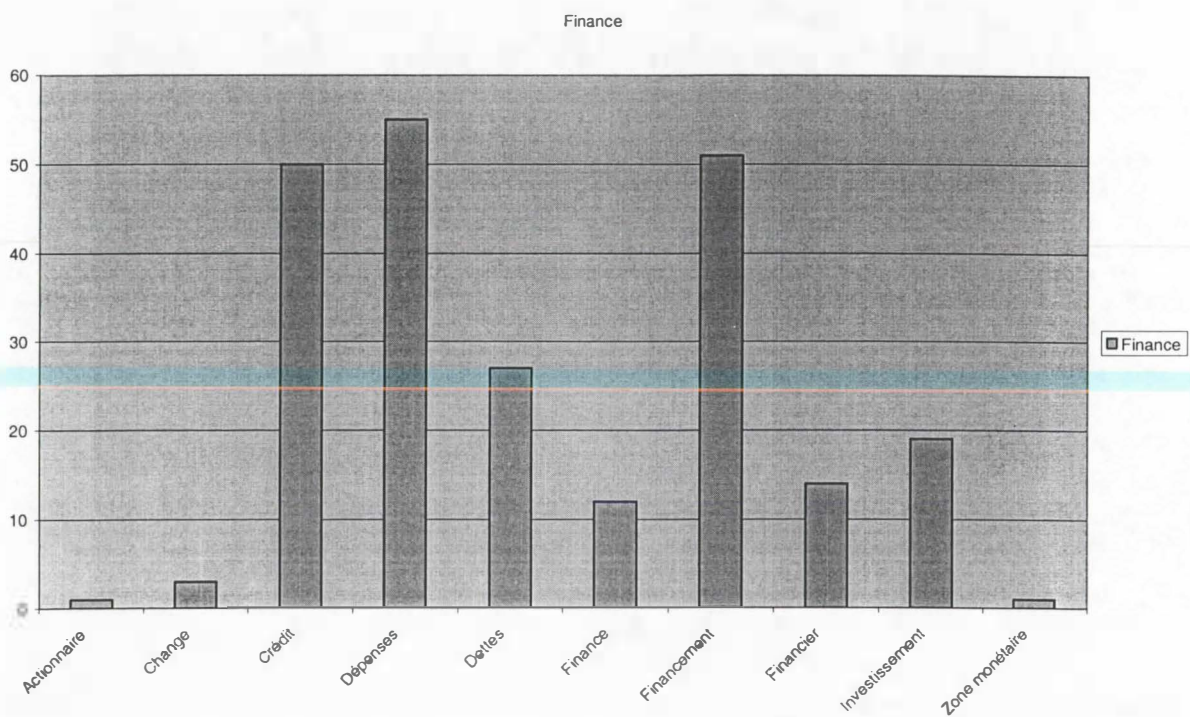


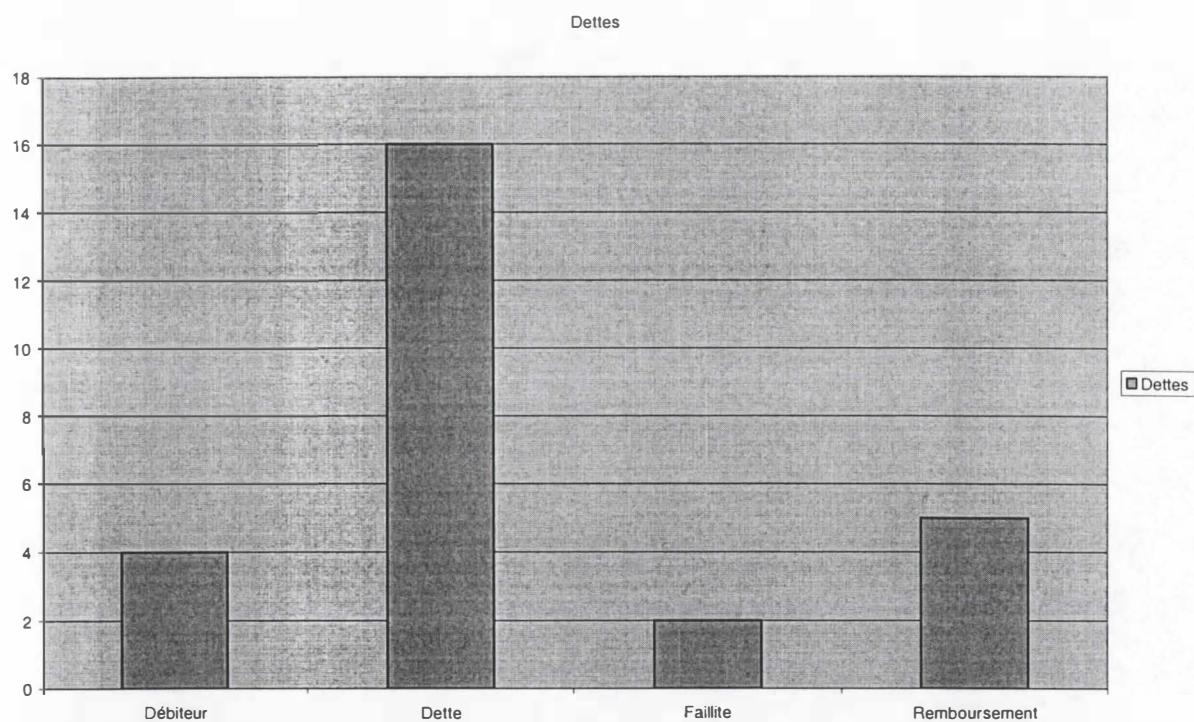
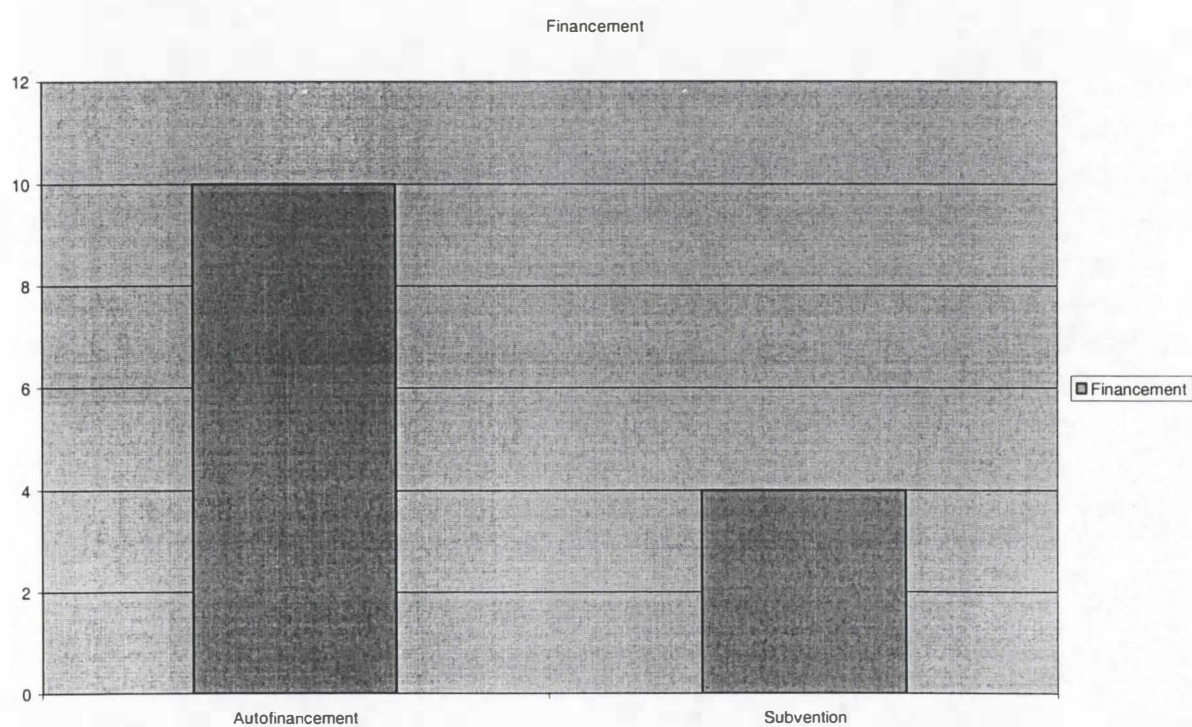
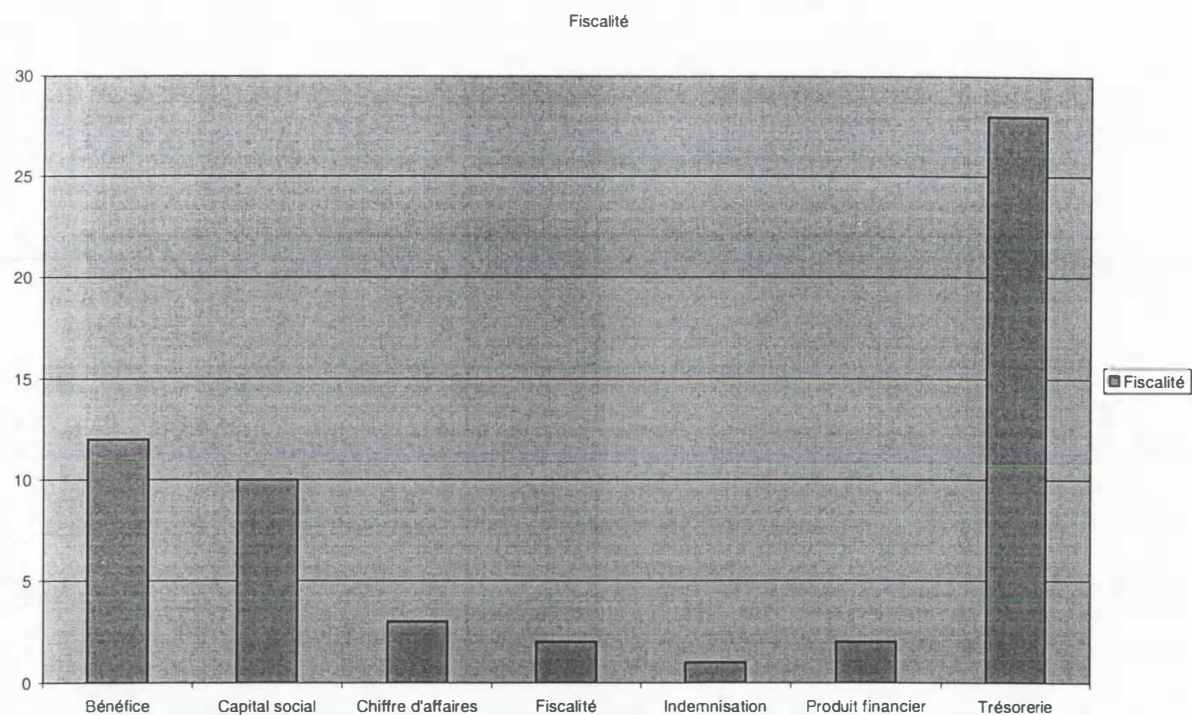
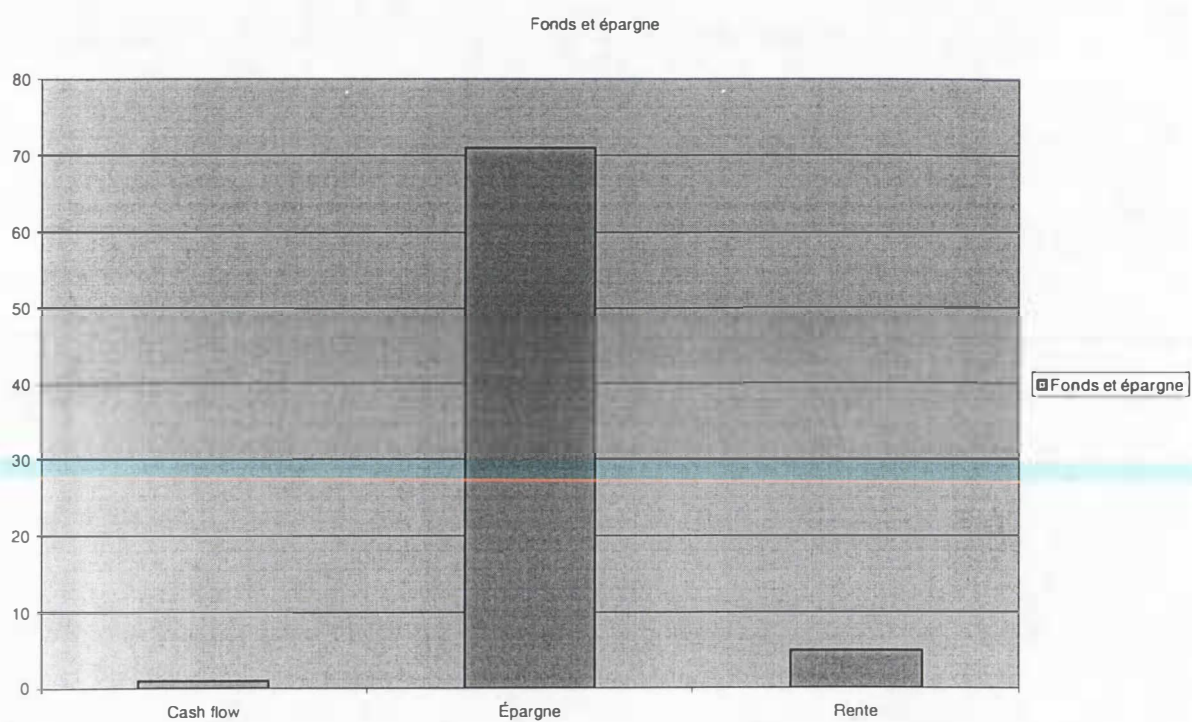
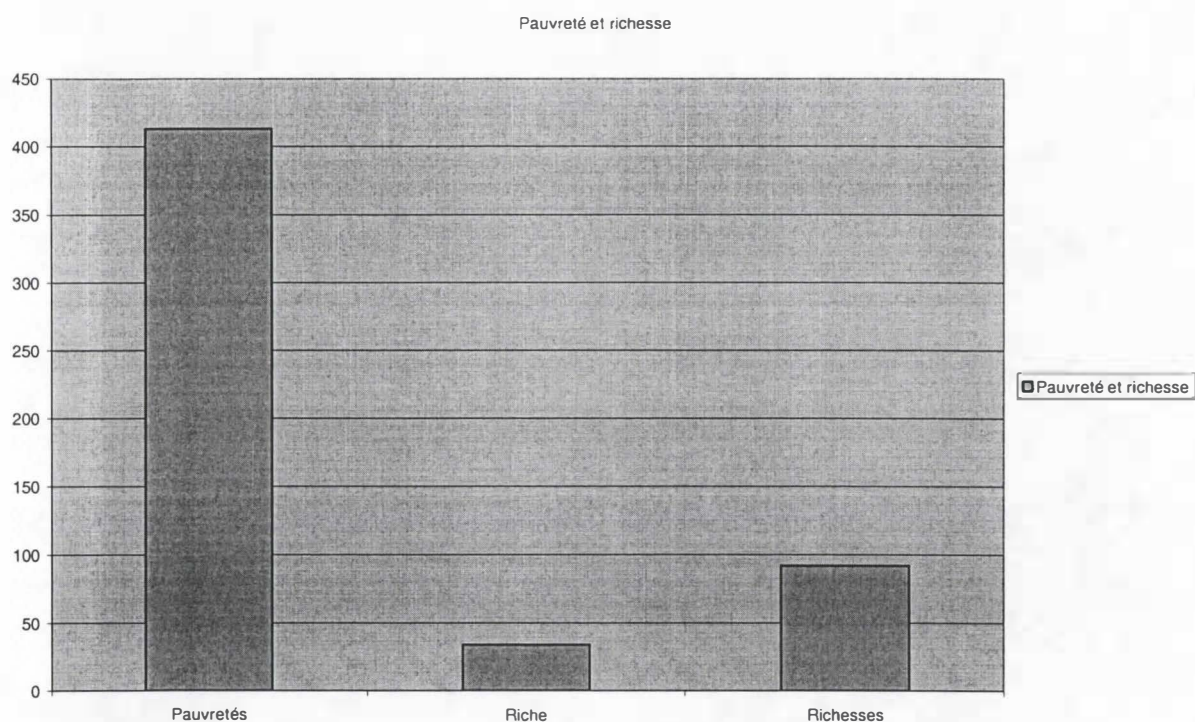
Figure 9 : Répartition des mots dans la sous-classe Dettes (27 mots)**Figure 10 : Répartition des mots dans la sous-classe Financement (14 mots)**

Figure 11 : Répartition des mots dans la sous-classe Fiscalité (58 mots)**Figure 12 :** Répartition des mots dans la sous-classe Fonds et épargne (77 mots)

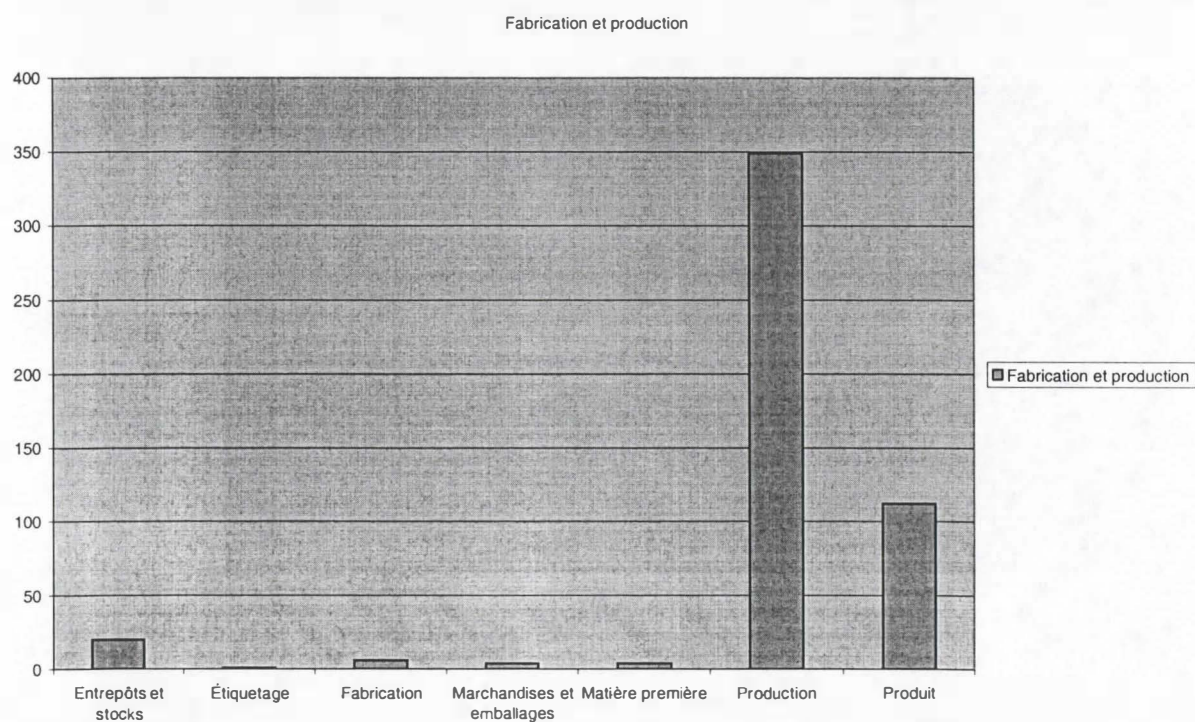
5.2.2.4. Pauvreté et richesse

Figure 13 : Répartition des mots dans la classe Pauvreté et richesse (539 mots)



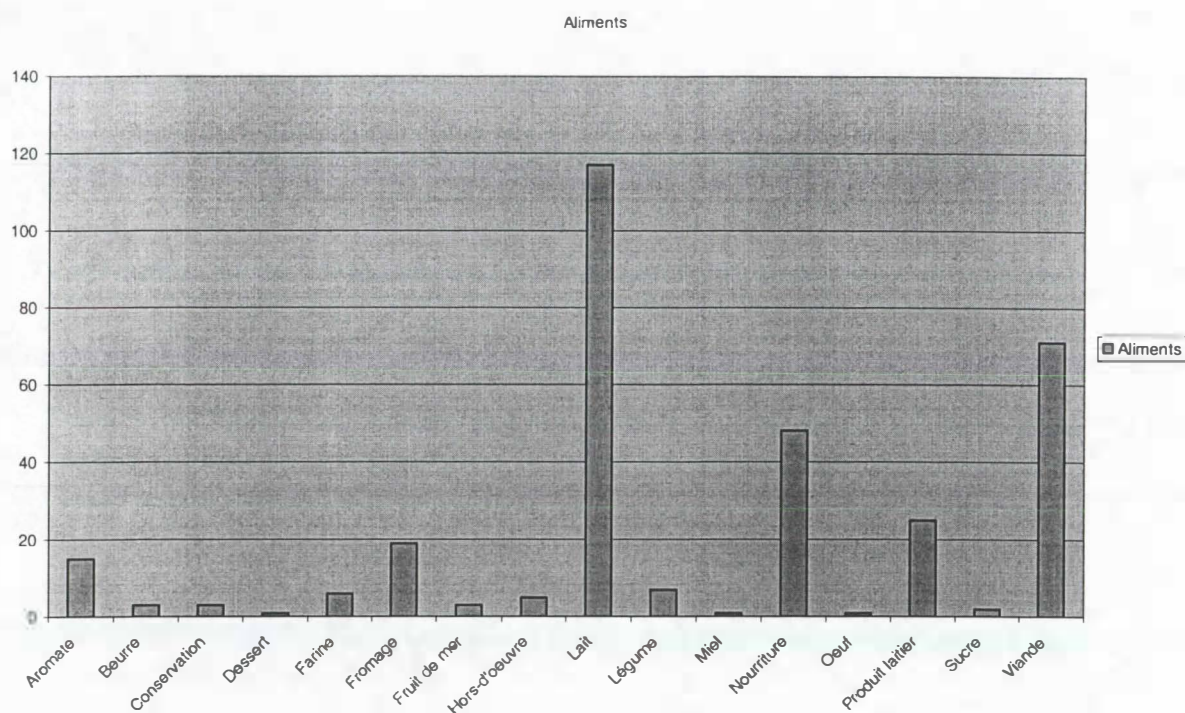
5.2.2.5. Fabrication et production

Figure 14 : Répartition des mots dans la classe Fabrication et production (496 mots)



5.2.2.6. Aliments

Figure 15 : Répartition des mots dans la classe Aliments (328 mots)



5.2.2.7. Politique et syndicalisme

Figure 16 : Répartition des mots dans la classe Politique et syndicalisme (278 mots)

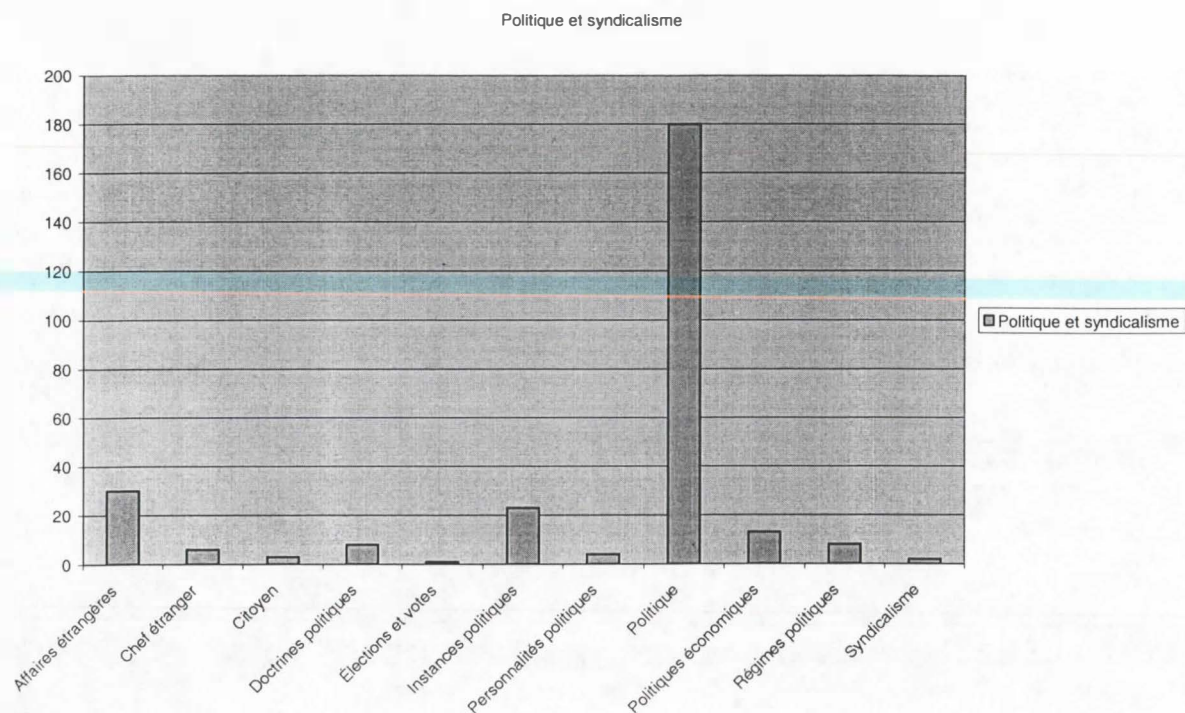


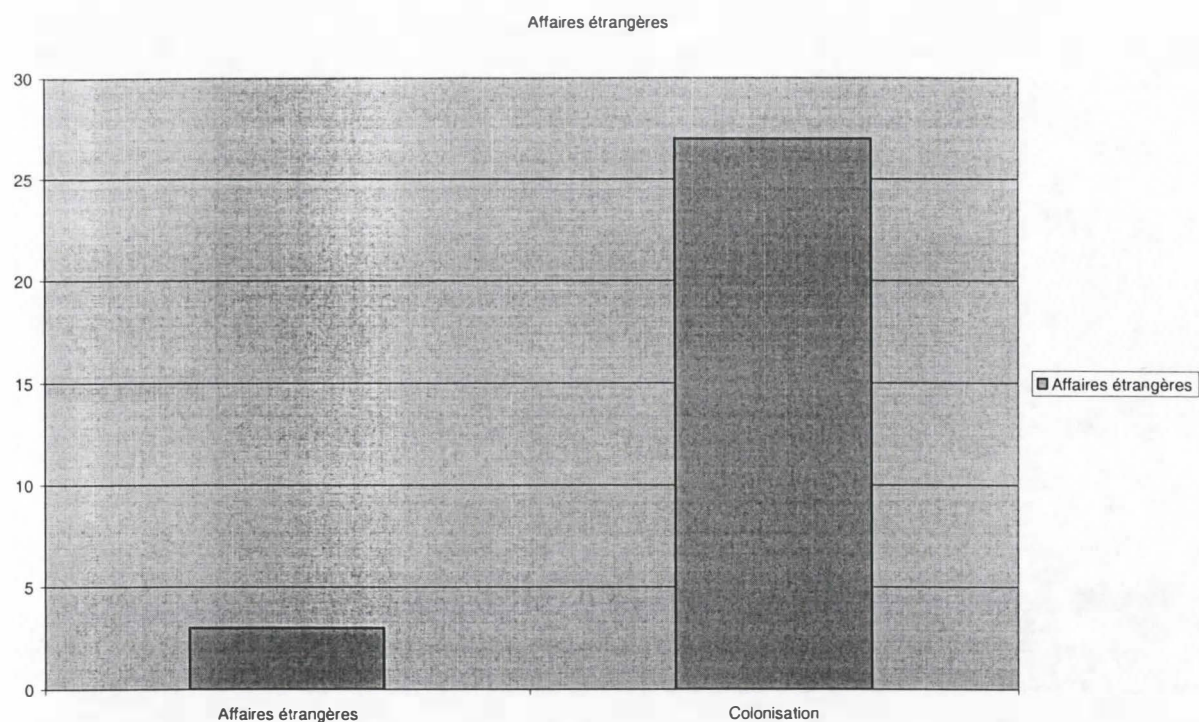
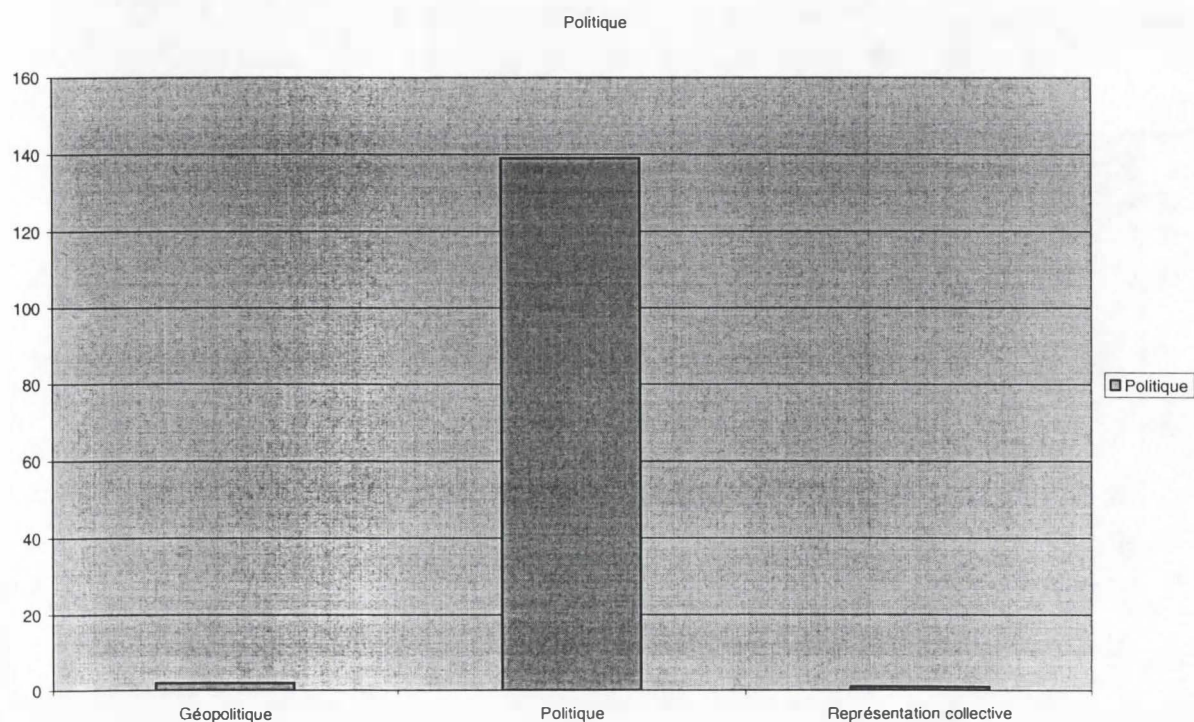
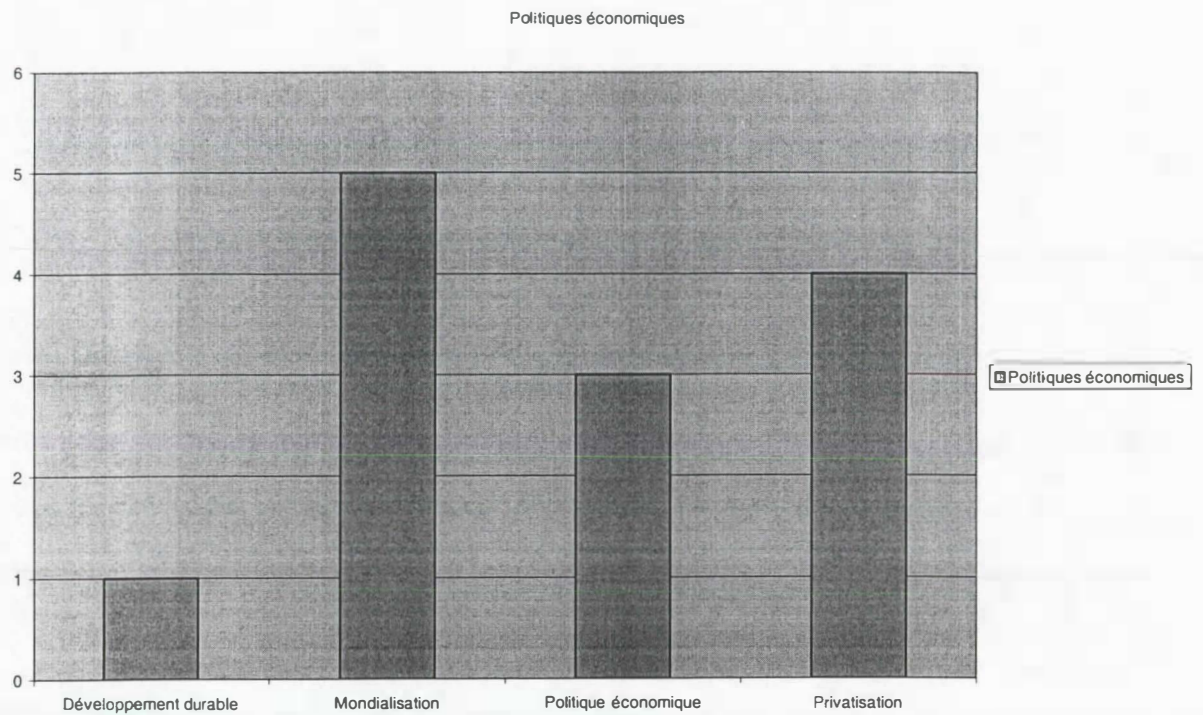
Figure 17 : Répartition des mots dans la sous-classe Affaires étrangères (30 mots)**Figure 18 : Répartition des mots dans la sous-classe Politiques (142 mots)**

Figure 19 : Répartition des mots dans la sous-classe Politiques économiques (13 mots)

5.2.2.8. Pays

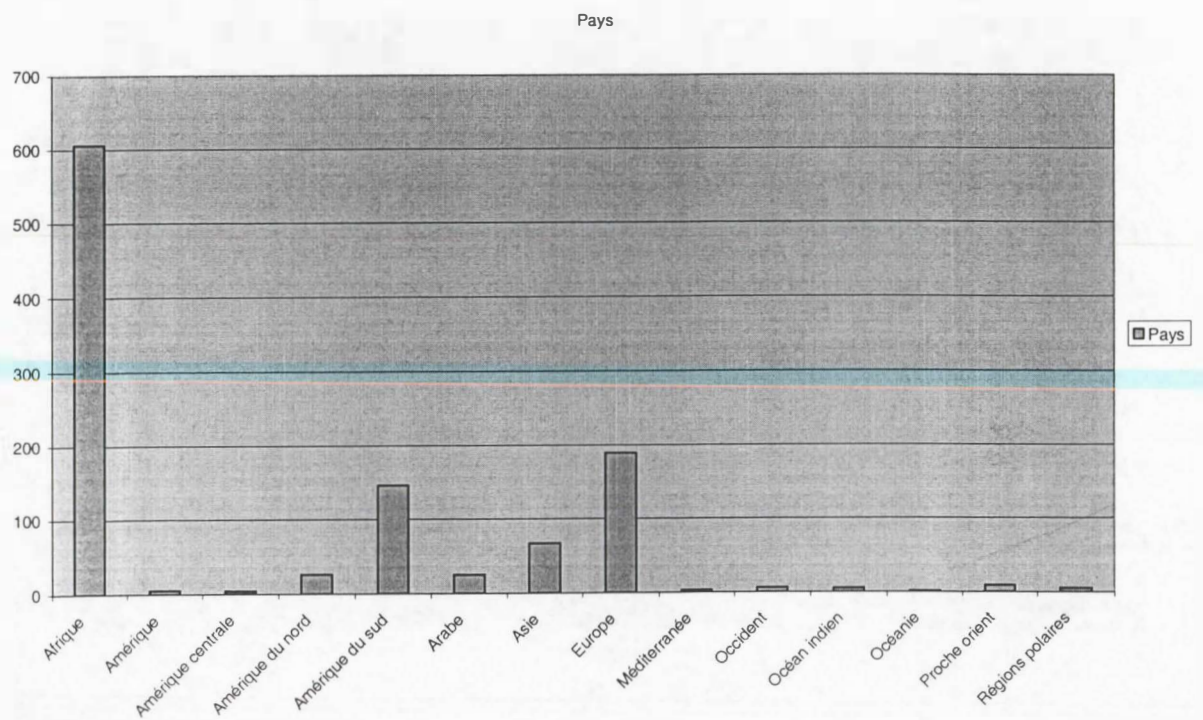
Figure 20 : Répartition des mots dans la classe Pays (1097 mots)

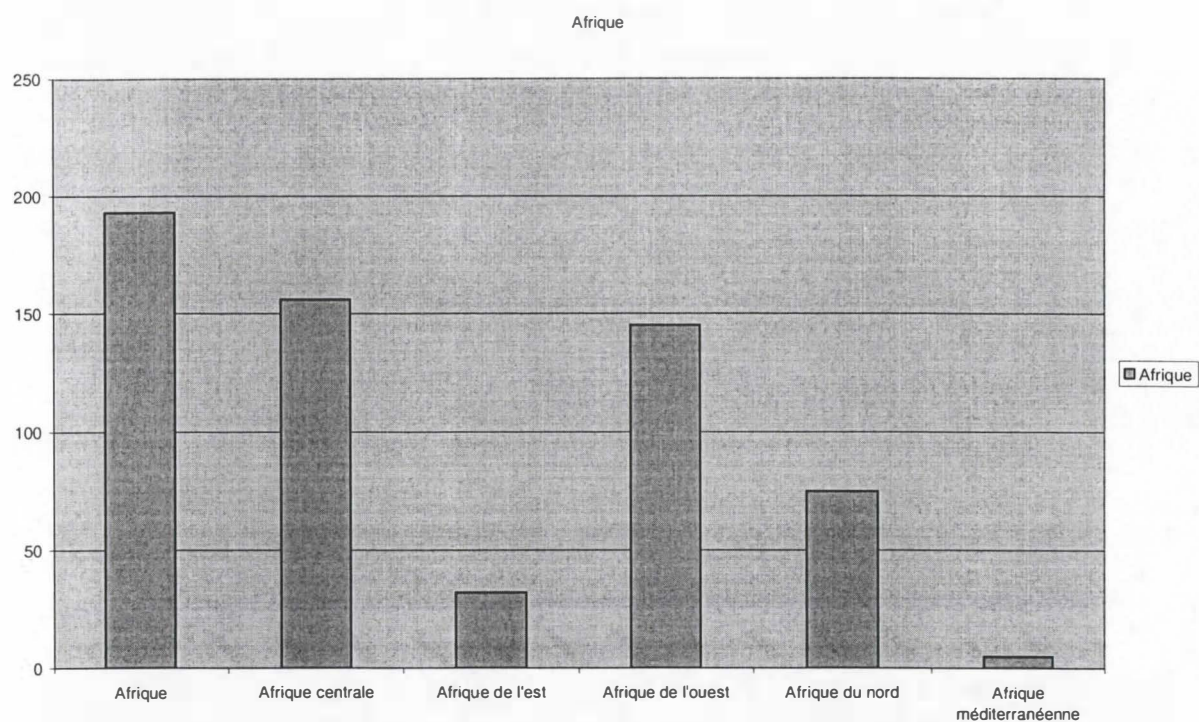
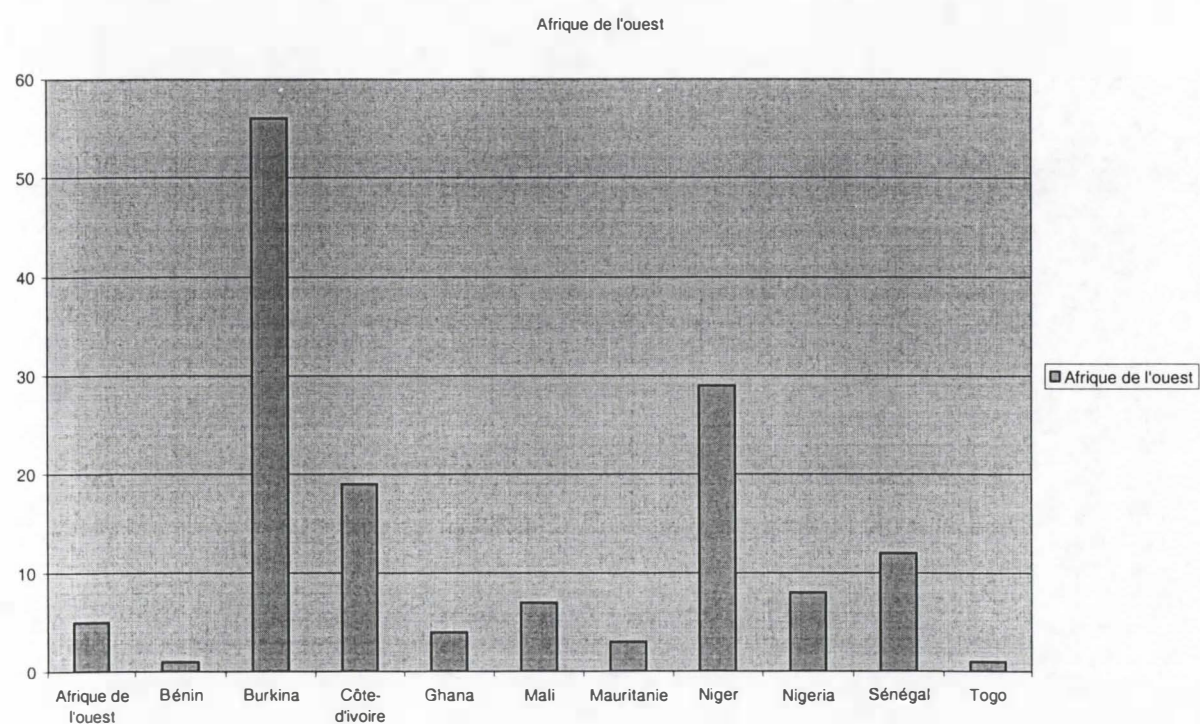
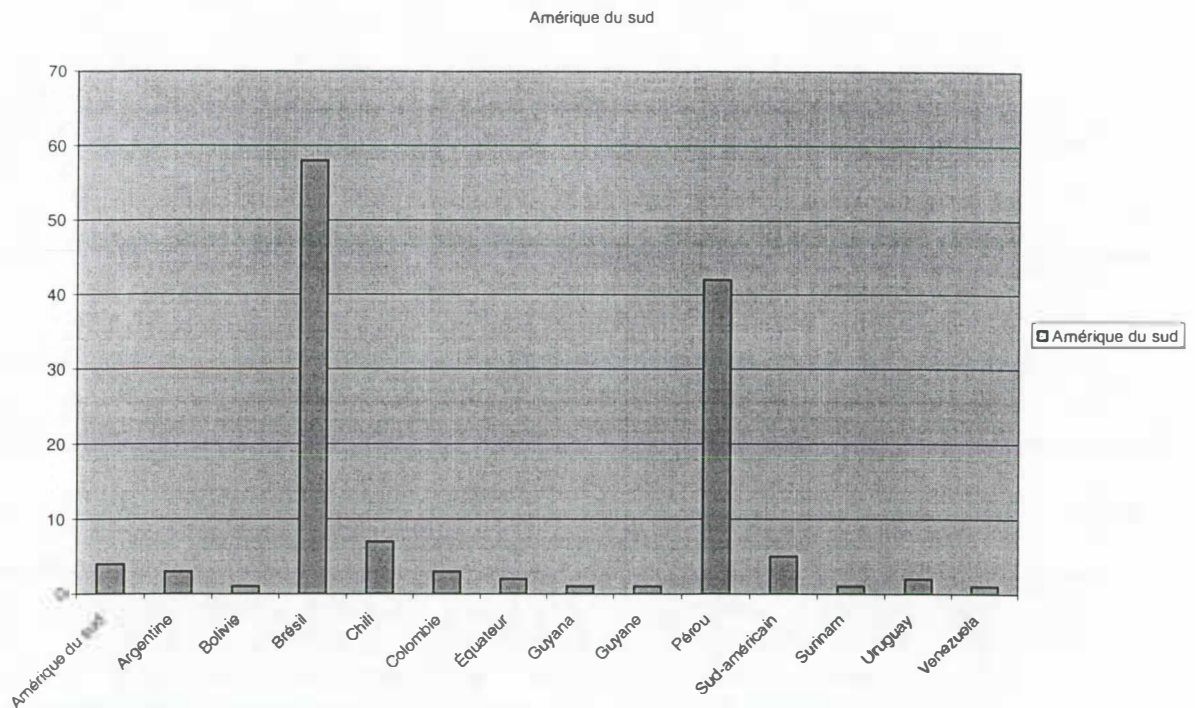
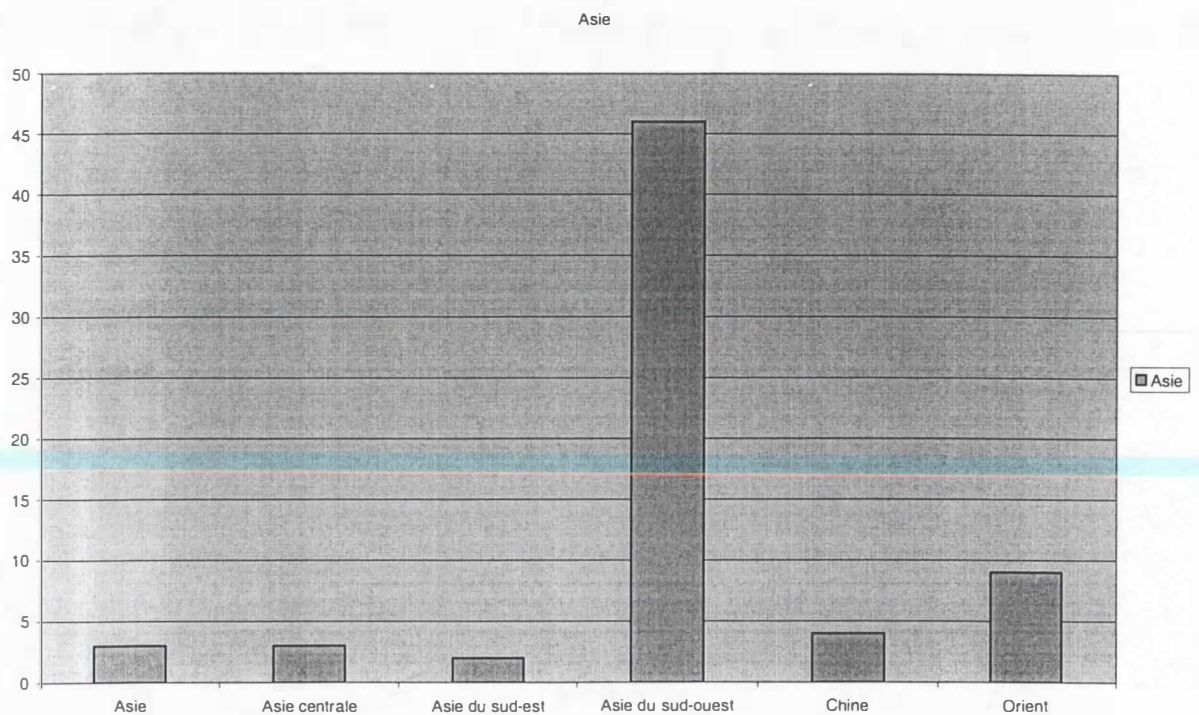
Figure 21 : Répartition des mots dans la sous-classe Afrique (606 mots)**Figure 22 :** Répartition des mots dans la sous-classe Afrique de l'Ouest (145 mots)

Figure 23 : Répartition des mots dans la sous-classe Amérique du Sud (146 mots)**Figure 24 : Répartition de mots dans la sous-classe Asie (67 mots)**

6. Synthèse des résultats

Les textes analysés correspondent aux contributions de différents auteurs, rédigées à l'occasion d'un atelier de réflexion sur l'élevage et la pauvreté rurale.

Compte tenu des thèmes de recherche annoncés, à savoir la pauvreté et les stratégies de réduction de la pauvreté, la sécurité alimentaire et l'accès aux services, le rôle de l'agriculture et de l'élevage dans la réduction de la pauvreté, le développement des pays du sud, il est intéressant d'examiner, à partir des résultats d'analyse sémantique de *Tropes Zoom*, en quels termes les différentes notions sont traduites dans les écrits.

L'un des thèmes prioritaire et traité par tous les auteurs (17 sur 17 documents) est **agriculture et élevage** (figures 4 à 6).

Les contributions parlent plus d'élevage (706 mots) que d'agriculture (418 mots).

Au sein de la classe Elevage, les concepts abordés sont peu diversifiés : on parle essentiellement d'éleveur (295 mots) ou de berger (103 mots).

Dans la catégorie **économie et commerce** (figures 1 à 3), notions couvertes par tous les documents, il est fréquemment question d'économie et de consommation (383 mots), de marché et de commerce (323 mots).

Les auteurs parlent moins de marketing et vente (251 mots) ou de prix (172 mots), rarement d'import-export (53 mots) et très rarement de client (8 mots).

Lorsque les auteurs traitent de la question marketing et vente, ils parlent surtout de vente (139 mots) et de commercialisation (91 mots), pratiquement jamais de marketing (1 mot).

De la même façon, au sein de la sous-classe marché et commerce, les concepts de fournisseur (1 mot), d'entrepreneurs (1 mot) et de marchand (8 mots) sont très rarement introduits.

La classe **pauvreté et richesse** (figure 13) est plus faiblement représentée (539 mots) bien qu'objet central de l'atelier dans lequel s'inscrivent les textes analysés. Les mots les plus fréquemment rencontrés sont pauvreté (413 mots), riche ou richesses (126 mots).

Les notions de tiers-monde et de pays en développement ne sont par contre quasiment jamais abordées (3 mots) par les différents auteurs.

Les sujets les plus fréquemment abordés dans la classe **argent et finance** (627 mots) (figures 7 à 12), sont ceux se rapportant aux finances (233 mots) comme les dépenses (55 mots) ou les crédits (50 mots), et ceux relevant des moyens de paiement (181 mots) comme les monnaies (76 mots).

Les auteurs parlent par contre peu de banques (31 mots), de budgets (13 mots), encore moins de comptabilité (3 mots). La notion de gaspillage n'est évoquée qu'une seule fois.

En termes de **fabrication et de production** (496 mots) (figure 14), l'accent est mis sur la production (349 mots) et plus particulièrement sur le producteur lui-même (134 mots).

Les produits (112 mots) sont plus faiblement évoqués.

Encore plus rarement abordés sont les concepts d'entrepôts et stocks (20 mots), de fabrication (6 mots), de matière première (4 mots), de marchandises ou d'emballages (4 mots) et d'étiquetage (1 mot).

Parmi les **aliments** (328 mots) (figure 15), ce sont le lait (117 mots) et la viande (71 mots) qui sont le plus fréquemment cités, suivis des produits laitiers (25 mots).

Les autres aliments d'origine animale tels que le fromage (19 mots), le beurre (3 mots), l'œuf (1 mot) ou le miel (1 mot) sont beaucoup plus rarement évoqués.

Il est à noter, que le concept fruit de mer reconnu trois fois par le scénario, est impropre dans le contexte étudié, s'agissant de tourteaux de coton. Une modification du scénario aurait

permis de lever cette confusion sémantique.

Enfin, le concept de **politique et syndicalisme** (278 mots) (figures 16 à 19) se résume à celui de politique au sens large (180 mots). Et, pour l'ensemble des communications, le concept plus spécifique de politique économique (13 mots) est peu utilisé, ni celui de géopolitique (2 mots).

Les notions de colonisation (27 mots), d'instances politiques (23 mots), de politiques économiques (13 mots) sont plus rarement évoquées et encore moins, celles de doctrine politique (8 mots), de régimes politiques (8 mots), de chef étranger - comme le sultan - (6 mots), de personnalités politiques (4 mots) et de citoyen (3 mots).

Quant à la géographie des lieux d'action ou d'écriture (les mentions d'adresses des auteurs n'ont pas été omises de l'analyse), elle est regroupée dans la classe **pays** (figures 20 à 24) riche de 1097 mots.

L'Afrique est largement représentée (606 mots) avec l'Afrique centrale (156 mots) et l'Afrique de l'Ouest (145 mots).

Plus faiblement traitées sont l'Afrique du nord (75 mots) et l'Afrique de l'Est (32 mots).

L'Afrique du Sud est absente des communications.

L'Amérique se réduit essentiellement à l'Amérique du Sud (146 mots) avec le Brésil (58 mots) et le Pérou (42 mots), pays les plus fréquemment cités.

L'Amérique centrale est très peu évoquée (4 mots) dans les différents écrits.

Les autres régions sont peu ou pas couvertes : l'Asie (67 mots), le Proche-Orient (9 mots), l'Océan indien (5 mots), l'Océanie (1 mot).

Il serait intéressant, au vu de ces grandes variations de localisation géographique des actions de recherche du Cirad sur le thème élevage et pauvreté, de justifier ces choix et l'exclusion des zones géographiques sus citées.

7. Conclusions

L'analyse de textes par *Zoom Tropes* repose sur des grammaires, des dictionnaires et des classifications sémantiques élaborées.

Sur la base de scénarios sémantiques, *Zoom Tropes* effectue plusieurs types d'analyse : statistique, probabiliste et cognitivo-discursive.

L'intérêt d'un tel outil réside dans la vitesse de lecture des documents et l'automatisation de l'analyse sémantique et statistique. Celle-ci assure une objectivité d'interprétation de textes.

Une analyse par *Tropes Zoom* n'invente rien. En ce sens, la lecture de ses résultats peut paraître décevante.

Elle permet toutefois de détecter rapidement et de visualiser automatiquement les fluctuations sémantiques entre documents ou à l'intérieur d'un texte.

La lecture de résultats d'analyse peut être ainsi révélatrice de certaines lacunes. Elle peut au contraire faire apparaître certaines convergences sémantiques ou des concepts émergents.

La présente étude, réalisée à titre expérimental pour une première approche sémantique de textes par *Zoom Tropes*, outil d'analyse nouvellement acquis par la Dic au Cirad, a été menée dans certaines limites :

- Le corpus de documents étudié n'a pas été modifié en vue de son analyse. L'intégralité des contenus des documents a été examinée. Un balisage préalable aurait permis d'écarter certaines données non significatives telles que l'adresse des auteurs ou la bibliographie en fin de texte.

- Le scénario original utilisé n'a pas été modifié pour intégrer des classes sémantiques nouvelles spécifiques au domaine couvert par les documents du corpus.
C'est ainsi que des classes comme « ressource », « travaux », « cirad », « indice », « lutte contre » n'ont pas été intégrées au scénario et de ce fait n'ont pas été l'objet d'une classification.
- Les relations entre les termes n'ont pas été traitées. Ces relations de cooccurrence indiquent quelles classes d'équivalents sont fréquemment reliées, c'est-à-dire rencontrées à l'intérieur d'une même proposition dans le texte analysé.
- De la même façon, les catégories de mots fréquemment utilisées n'ont pas été étudiées. L'examen des catégories de mots telles que verbes, connecteurs, pronoms personnels, modalisations, adjectifs, permet de mettre en évidence la mécanique d'un texte et ses rouages : Le texte étudié est-il plus dans l'argumentation, le doute ou l'affirmation ? Quel type de lien unit les concepts entre eux?

Malgré ces limitations, la lecture des premiers résultats issus de l'analyse sémantique des communications de l'atelier-recherche du Cirad sur « Elevage et pauvreté » tenu en 2003, pourra faciliter l'identification, par le réseau de recherche « économie et politiques d'élevage », de thèmes prioritaires en vue de l'élaboration son futur programme de recherches.

Bien que les concepts les plus partagés par les documents du corpus étudié s'inscrivent largement dans les thématiques annoncées, comme l'agriculture et l'élevage, l'économie et le commerce, l'argent et les finances, les aliments, les animaux ou la médecine et la santé, l'analyse sémantique des textes avec *Tropes Zoom* fait apparaître de grandes variations dans la représentation, au sein de ces classes génériques, de concepts plus spécifiques.

C'est ainsi qu'en économie, les questions inhérentes à la macro-économie et la micro-économie, aux phénomènes d'inflation et de déflation, d'exportation, de marketing, et d'entrepreneuriat sont faiblement abordés.

En politique, la géopolitique et, à un degré moindre, la politique économique, sont absentes des débats.

Les animaux étudiés sont essentiellement les bovidés et, moins fréquemment les ovins. Il est très peu question de cheval, de porc, ou des camélidé.

Les questions de médecine et santé se résument à la notion de santé et de maladies. La diététique ou la pharmacologie sont rarement l'objet de réflexion.

Parmi les aliments, les produits laitiers tels que le beurre ou le fromage ne sont pratiquement pas étudiés.

Par ailleurs, le concept de méthodologie est plus faiblement représenté que le concept de technologie ou de techniques.

Les auteurs débattent plus qu'ils n'argumentent.

Enfin, dans l'ensemble du corpus étudié, les causes sont moins souvent traitées que les effets.

Un examen approfondi par les auteurs eux-mêmes ou des spécialistes du domaine, de l'ensemble des résultats obtenus en nombre de documents (analyse *Zoom*) et en nombre de mots (analyse *Tropes*) devrait permettre d'expliquer plus précisément et de justifier les variations sémantiques observées comme la faible représentation de thèmes spécifiques au domaine, ou l'absence de concepts relevant pourtant du champ d'investigation annoncé.

8. Annexes

8.1. Annexe 1 : Liste des fichiers et titres des documents analysés

Nom de fichier Word (.doc)	Titre des communications
ARDITI COMMERCE BETAIL.DOC	Les conséquences de la paupérisation des éleveurs Peuls sur l'organisation des réseaux de commercialisation du bétail et de la viande en RCA
ARDITI PAUPERISATION.DOC	La paupérisation des éleveurs peuls de RCA
BOUCHER CAPABILITIES.DOC	La paupérisation des éleveurs peuls de RCA
BOURGEOT PAUVRETE.DOC	Pauvres, protection et dynamiques pastorales au sahel
CLANET ELEVEURS PAUVRES.DOC	Des éleveurs pauvres globalement, mais riches localement
DUTEURTRE LAIT DES PAUVRES.DOC	Lait des pauvres, lait des riches : l'impact des politiques de libéralisation sur l'accès aux marchés des éleveurs pauvres en Afrique
DUTEURTRE MICRO-ECONOMIE.DOC	Elevage et dynamique de la pauvreté : L'approche micro-économique
DUTILLY COOPERATION.DOC	Coopération pour la GRN : Impact sur le niveau, la diversification et l'inégalité des revenus au Sahel.
FAYE PAUVRETE ET INEGALITES.DOC	Pauvreté et solidarité chez les peuples pastoraux
HASSANPAUVRETE ET ELEVAGE DANS LA STEPPE NORD.DOC	Pauvreté et élevage dans la steppe nord-africaine
LE MASSON PRODUIRE ET BIEN VENDRE.DOC	Produire et bien vendre le bétail : survie des pasteurs et dynamique du troupeau
RENARD FILIERES DE COMMERCIALISATION.DOC	Filières de commercialisation et pauvreté : le cas des produits animaux au Burkina Faso
RIBIERGRAWITZLE THEME DE LA PAUVRETE DANS LES TRAVAUX DU CIRAD.DOC	Le thème de la pauvreté dans les travaux du CIRAD. Une parole novatrice dans les débats ?
TACHER RISQUE ET PAUVRETE.DOC	Note sur le risque et la pauvreté chez les éleveurs
TOURRAND ELEVAGE ET PAUVRETE.DOC	L'élevage bovin dans la réduction de la pauvreté de l'agriculture familiale en Amazonie brésilienne
VALL SEUILS DE PAUVRETE.DOC	Seuils de pauvreté et projets d'équipement agricole : Une étude de cas sur 4 exploitations cotonnières du Nord-Cameroun
VERONIQUEALARYRISQUE_INDE.DOC	L'élevage dans la gestion des risques par les producteurs du Nord de l'Inde

8.2. Annexe 2 : Liste des figures

Figure 1 : Répartition des mots dans la classe Economie et commerce (1275 mots)	11
Figure 2 : Répartition des mots dans la sous-classe Marché et commerce (323 mots).....	11
Figure 3 : Répartition des mots dans la sous-classe Marketing et vente (251 mots).....	12
Figure 4 : Répartition des mots dans la classe Agriculture et élevage (1265 mots)	12
Figure 5 : Répartition des mots dans la sous-classe Agriculture (418 mots).....	13
Figure 6 : Répartition des mots dans la sous-classe Elevage (706 mots)	13
Figure 7 : Répartition des mots dans la classe Argent et finance (627 mots).....	14
Figure 8 : Répartition des mots dans la sous-classe Finance (233 mots)	14
Figure 9 : Répartition des mots dans la sous-classe Dettes (27 mots)	15
Figure 10 : Répartition des mots dans la sous-classe Financement (14 mots)	15
Figure 11 : Répartition des mots dans la sous-classe Fiscalité (58 mots).....	16
Figure 12 : Répartition des mots dans la sous-classe Fonds et épargne (77 mots)	16
Figure 13 : Répartition des mots dans la classe Pauvreté et richesse (539 mots)	17
Figure 14 : Répartition des mots dans la classe Fabrication et production (496 mots).....	17
Figure 15 : Répartition des mots dans la classe Aliments (328 mots).....	18
Figure 16 : Répartition des mots dans la classe Politique et syndicalisme (278 mots).....	18
Figure 17 : Répartition des mots dans la sous-classe Affaires étrangères (30 mots).....	19
Figure 18 : Répartition des mots dans la sous-classe Politiques (142 mots)	19
Figure 19 : Répartition des mots dans la sous-classe Politiques économiques (13 mots) ...	20
Figure 20 : Répartition des mots dans la classe Pays (1097 mots).....	20
Figure 21 : Répartition des mots dans la sous-classe Afrique (606 mots).....	21
Figure 22 : Répartition des mots dans la sous-classe Afrique de l'Ouest (145 mots).....	21
Figure 23 : Répartition des mots dans la sous-classe Amérique du Sud (146 mots)	22
Figure 24 : Répartition de mots dans la sous-classe Asie (67 mots).....	22

8.3. Annexe 3 : Liste des tableaux

Tableau 1 : Classes sémantiques présentant, en leur sein, des variations importantes de thèmes, exprimées en nombre de documents concernés.....	6
Tableau 2 : Univers de référence 1, univers de référence 2 et références utilisées les plus fréquentes	10